Web application log monitoring with SPLUNK

Splunk is a powerful platform for searching, monitoring, and analyzing machine-generated data. It facilitates operational intelligence and is commonly used for log management and data visualization in IT environments.

Overview: The project is about setting up a simple website and monitoring the traffic in Splunk.

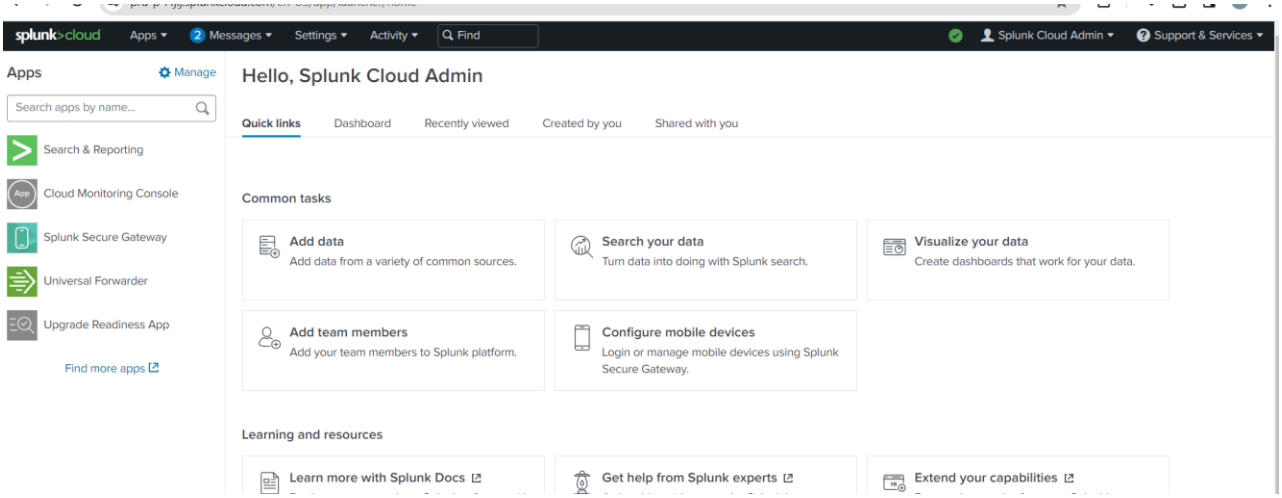
Two an EC2 machine are created with AWS

Machine 1- To Host a simple website, Install Splunk forwarder

Machine 2 – Install splunk, configure receiver port.

With this we can Monitor the website log in Splunk tool(Machine2).

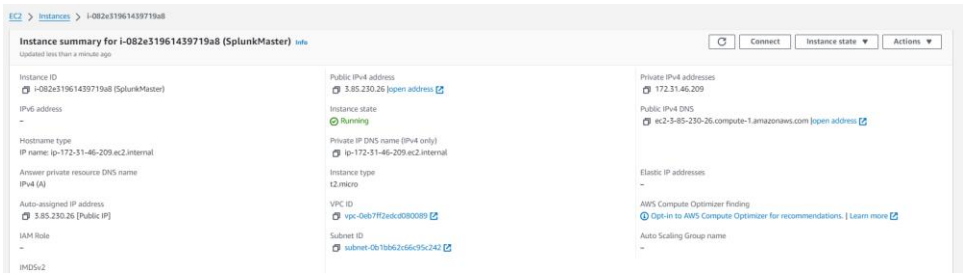
1. Create a splunk enterprise free account. Login to the account.



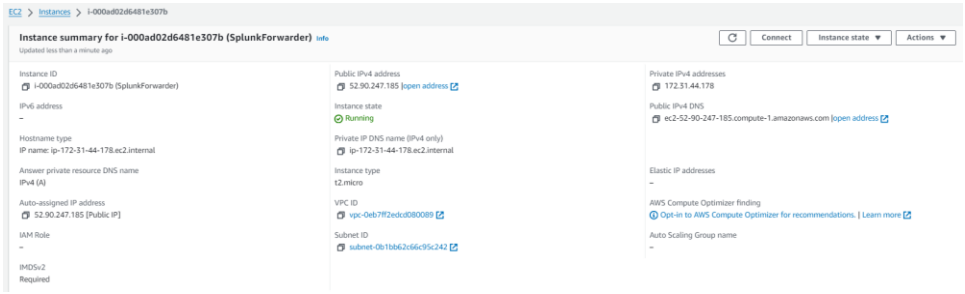
2. Create two EC2 machines in AWS, one to host the website and one to install SplunkMaster. Make sure to create a key pair and have it locally.

Select Redhat , t3.Medium for Splunk Master and t2.small for to host website, create new keypair if you don't have one or use the existing one, Allow ssh, http, https traffic from your ip or anywhere.

Splunk Master



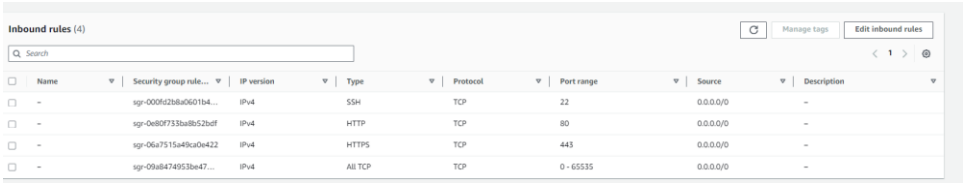
Splunk Forwarder



3. Change the key permission to 400(Only Read),

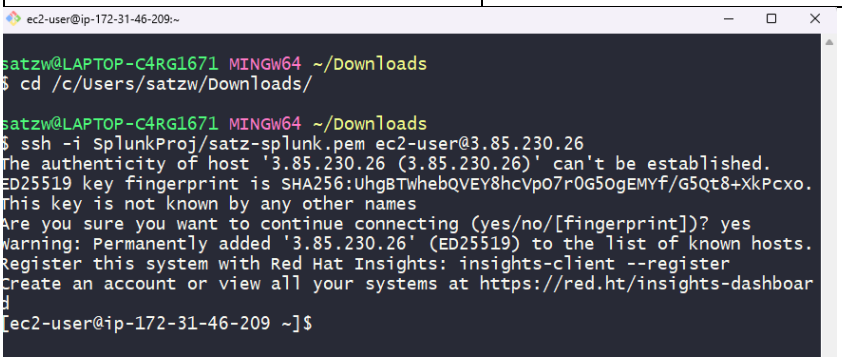
4. Allow all inbound TCP traffic from the public for the security tagged to both the Ec2 Instance,

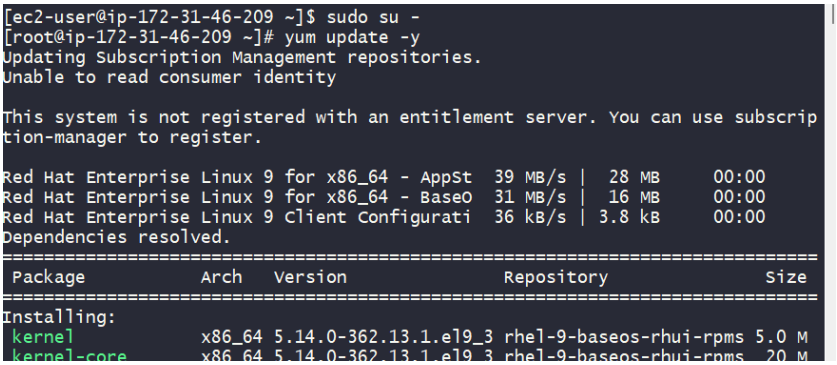
This is not recommended; we should allow only specific ip which is really required, just for demo, I have done this.



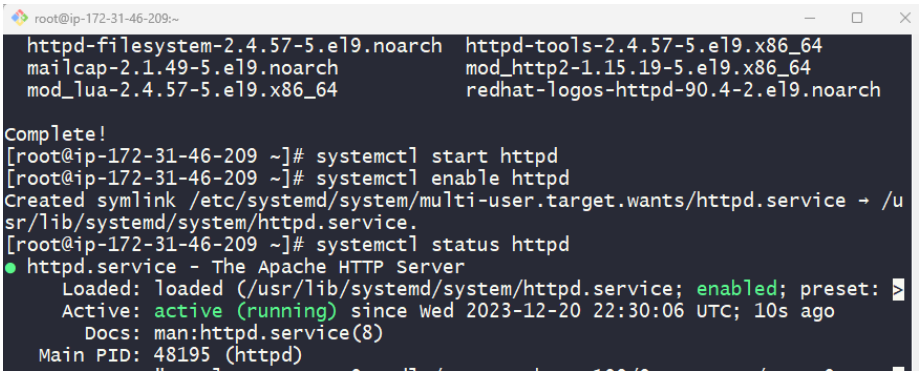
5. Open git bash and Login to Splunk Master EC2 instance with the Publicip and the Key we created.

cd /c/Users/satzw/Downloads/	Navigate to your local directory which has key file
ssh -i SplunkProj/satz-splunk.pem ec2-user@3.85.230.26	Ssh into the splunk master instance
sudo su -	Switch to root user
yum update -y	Update the yum package manager
yum install httpd	To install httpd web server
systemctl start httpd	To start the httpd server
systemctl enable httpd	To enable the httpd server
systemctl status httpd	To check the status of the httpd server
yum install wget -y	It's a is a utility for downloading files from the web

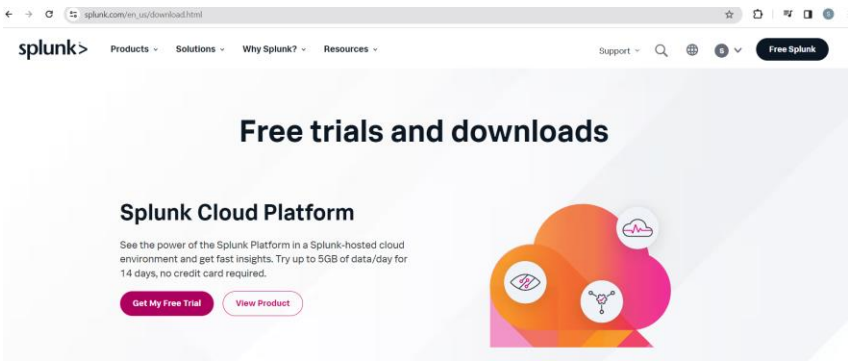


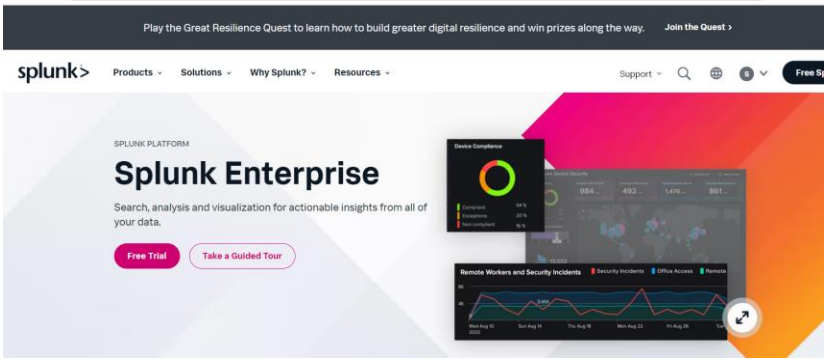


Below means the httpd(apache) webservice is active and running.

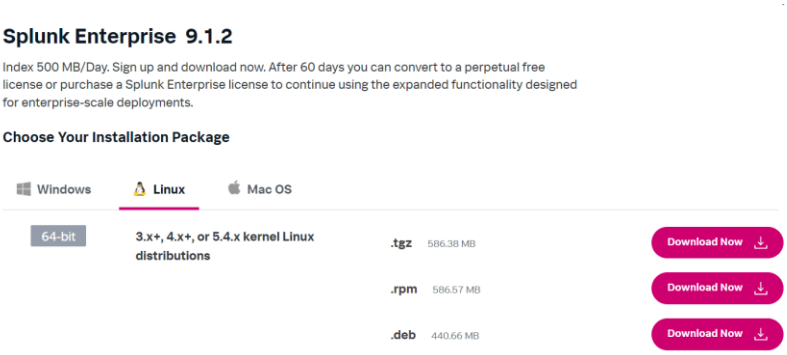


6. Create an account in splunk and get free trail of enterprise version to download splunk locally

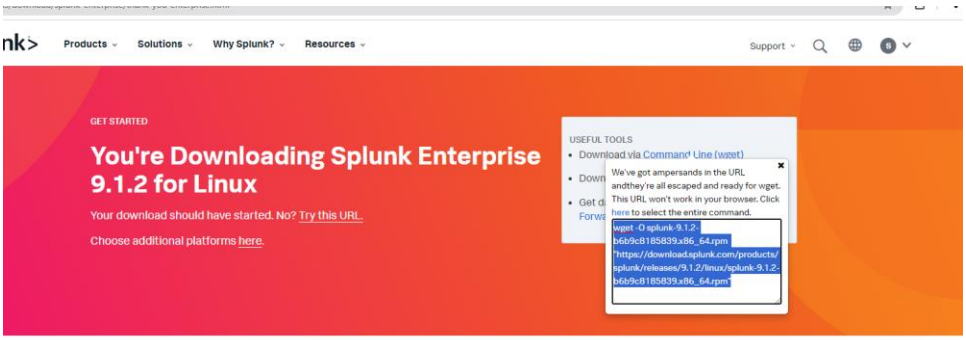




Click on free trail and download the .rpm package under Linux (since we use redhat OS)

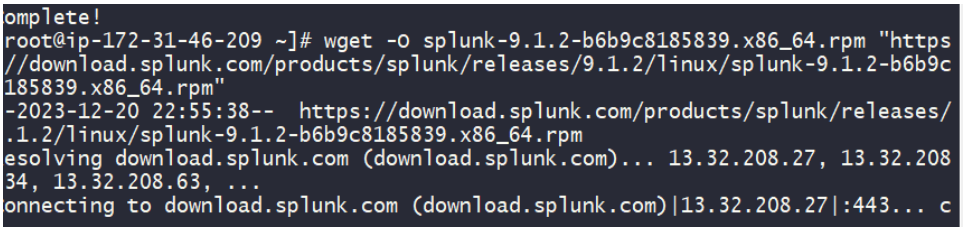


Once your download started, click on the command line wget link, copy the url, we can use it to directly download on our splunk master machine.



7. Now download the splunk package wth wget

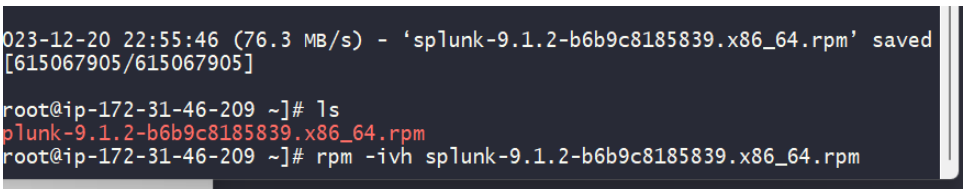
wget -O splunk-9.1.2-b6b9c8185839.x86_64.rpm
https://download.splunk.com/products/splunk/releases/9.1.2/linux/splunk-9.1.2-b6b9c8185839.x86_64.rpm



8. Install the package with rpm

rpm -ivh splunk-9.1.2-b6b9c8185839.x86_64.rpm

I – install, v- verbose, h-hash mark indicates the progress of installation

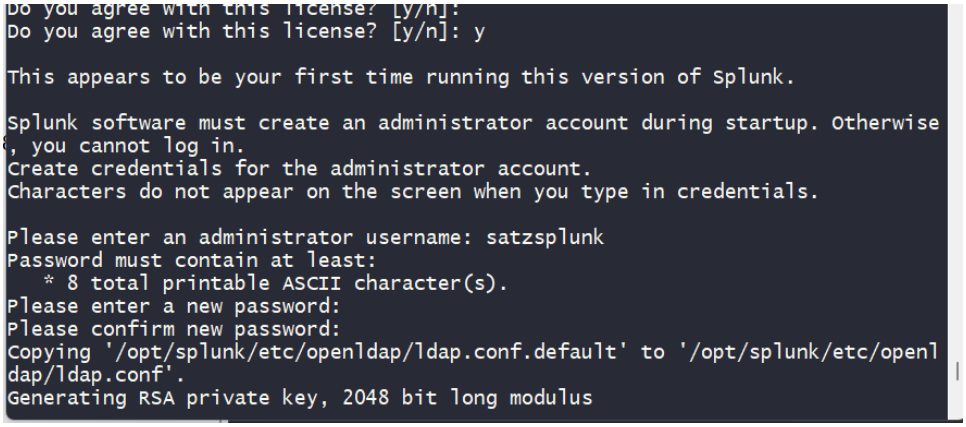


9. Get into the bin directory and start splunk

```
cd /opt/splunk/bin/

./splunk start --accept-license
```

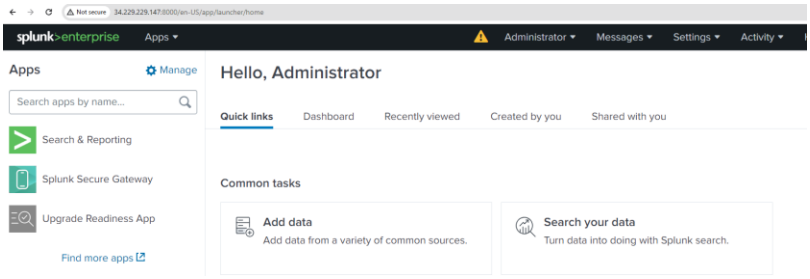
Enter Admin UserName and Password, now wait for splunk to start



10. Get the public ip of your EC2 Splunk master instance, open with http and port is 8000

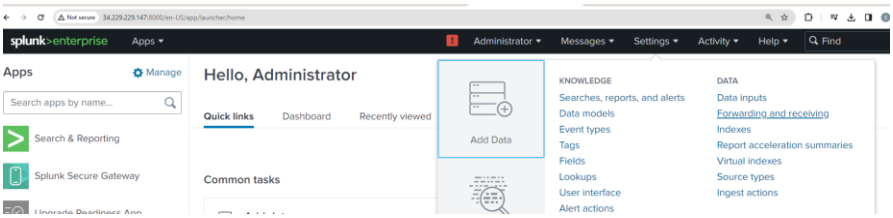
http://34.229.229.147:8000/ like this, replace your public ip

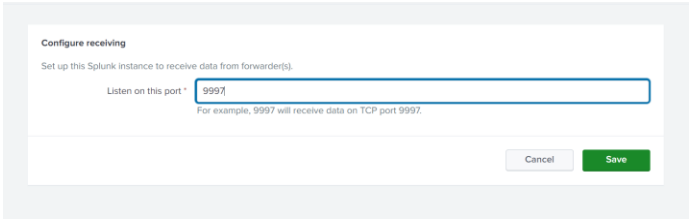
Login with username and pwd you gave earlier



11. Go to settings, under DATA select forwarding and receiving.

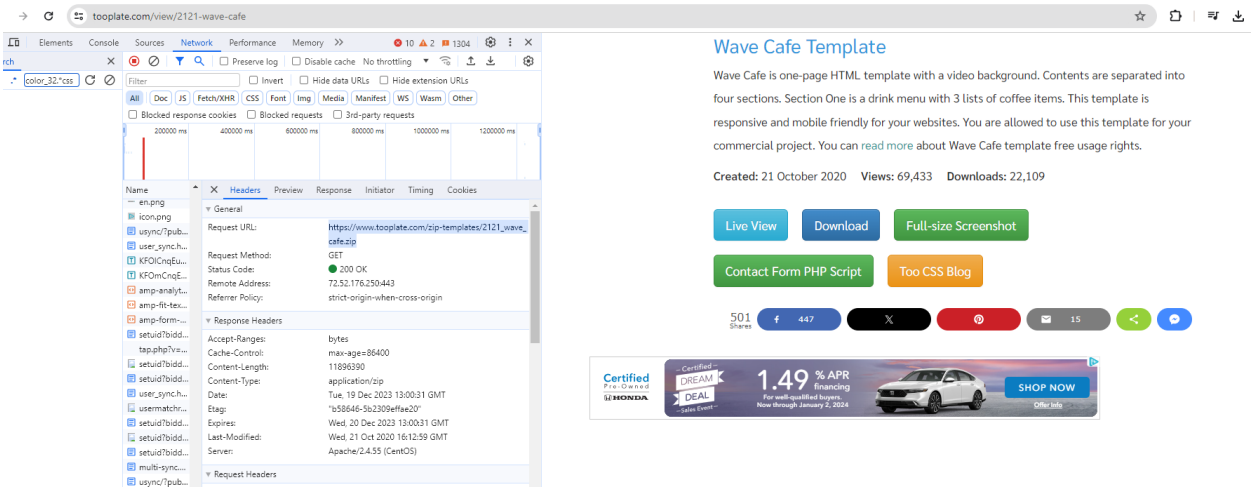
Click configure receiving, add new server port 9997 and save it,This means our master cluster is going to receive data on 9997 port from the other EC2 instance(which we will be hosting a website and install Splunk Forwarder.





12. It's Part-2 now, let's configure our website in the second EC2 instance.

Get the html package for a website from toplate.com. Goto tooplate.com, choose the html template you like, scroll down to see the download icon.Now click f12, which will open developer window on side (select Network tab, then select headers). Copy the url.



Wave cafe HTML: https://www.tooplate.com/zip-templates/2121_wave_cafe.zip

13. Ssh into the Second Ec2 Instance (Splunk 2) as explained on step 5(refer the table), install httpd and enable it.

Additionally install unzip package, this is required to unzip the website package files we are about to download

```
yum install unzip -y
```

14. Now Download the web application package to any directory

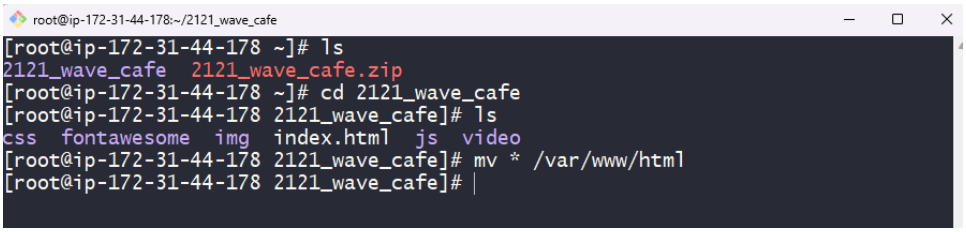
```
Complete!
[root@ip-172-31-44-178 ~]# wget https://www.tooplate.com/zip-templates/2121_wave
_cafe.zip
--2023-12-20 23:45:10--  https://www.tooplate.com/zip-templates/2121_wave_cafe.z
ip
Resolving www.tooplate.com (www.tooplate.com)... 72.52.176.250
Connecting to www.tooplate.com (www.tooplate.com)|72.52.176.250|:443... connecte
d.
HTTP request sent, awaiting response... 200 OK
Length: 11896390 (11M) [application/zip]
Saving to: '2121_wave_cafe.zip'

2121_wave_cafe.zip  100%[===================>]  11.34M  27.6MB/s    in 0.4s
```

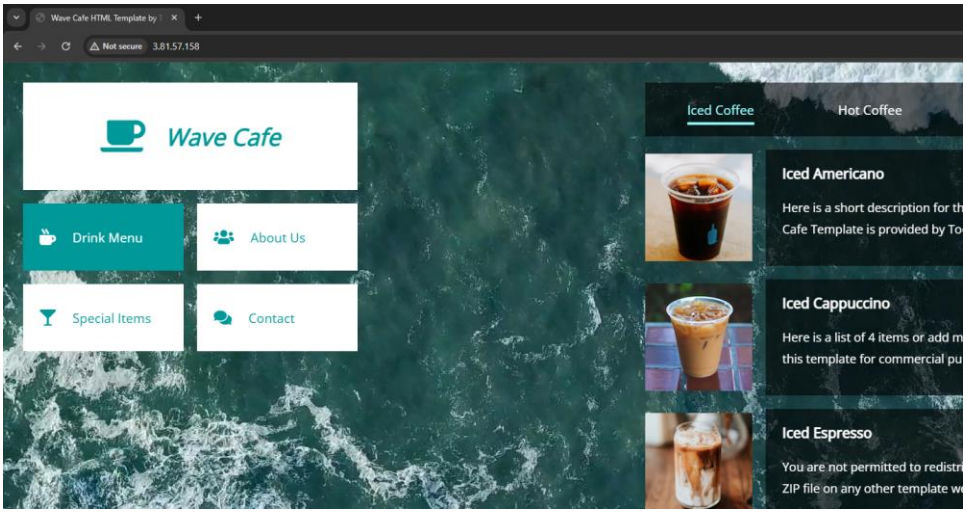
Unzip the file

```
[root@ip-172-31-44-178 ~]# unzip 2121_wave_cafe.zip
Archive:  2121_wave_cafe.zip
   creating: 2121_wave_cafe/
   creating: 2121_wave_cafe/css/
  inflating: 2121_wave_cafe/css/tooplate-wave-cafe.css
   creating: 2121_wave_cafe/fontawesome/
   creating: 2121_wave_cafe/fontawesome/css/
```

Get into the café directory, Move all the files into html directory. This is required since our Apache webserver will check the index.html file on this specific directory



15. The website is accessed with EC2 instance Public Ip. Its http not https

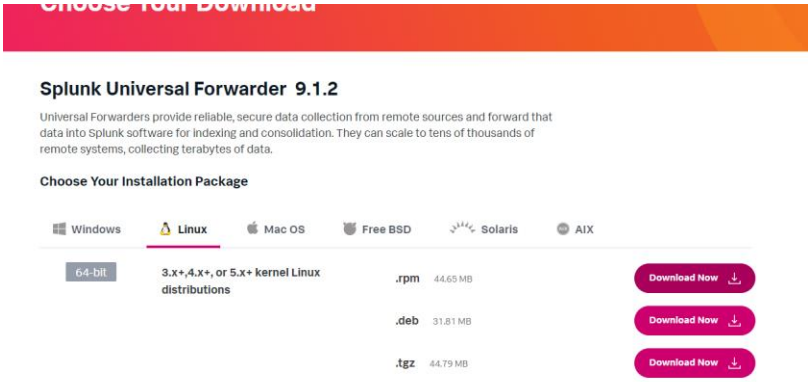


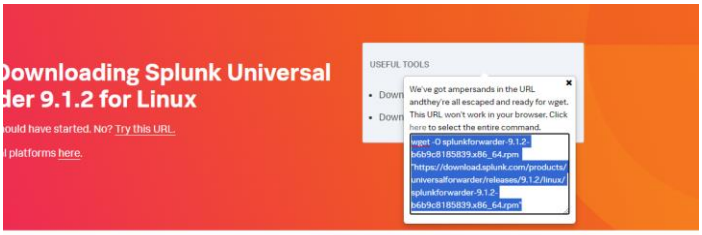
In case you are unable to access the site you hosted with Public ip in http

Troubleshoot ideas – Check https://raw.githubusercontent.com/satzwebio/Splunk_On_EC2_Website_Log/main/Troubleshoot_Steps.txt

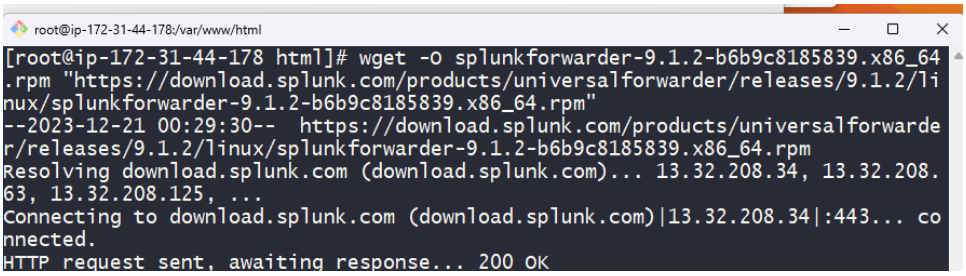
16. Let's install Splunk forwarder, which is required to forward the logs from the current EC2 machine (Machine 2- The website hosted) to Splunk master.

Search for 'splunk universal forwarder installation on linux' in google.Download the .rpm version, which means gather the wget url as like Step 6





wget -O splunkforwarder-9.1.2-b6b9c8185839.x86_64.rpm
https://download.splunk.com/products/universalforwarder/releases/9.1.2/linux/splunkforwarder-9.1.2-b6b9c8185839.x86_64.rpm



Install splunkforwarder we downloaded with rpm command

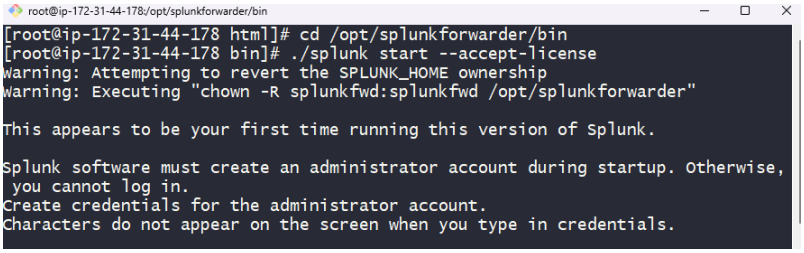
```
rpm -ivh splunkforwarder-8.0.2-a7878.rpm
```

Get into the appropriate directory and start the Splunk

```
cd /opt/splunkforwarder/bin
```

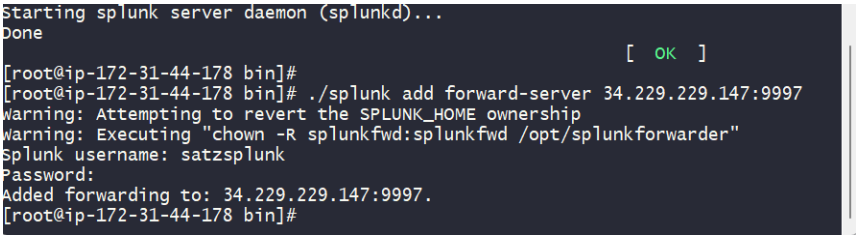
```
./splunk start --accept-license
```

enter Admin user, pwd



17. Add splunk master ip as the destination to send logs I.e. configure forwarder and the port should be 9997. Fyi - On Step 11, we mentioned the Master is going to receive traffic on 9997

```
./splunk add forward-server 34.229.229.147:9997
```



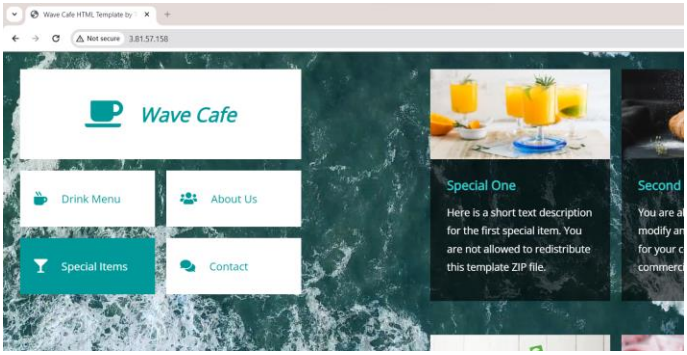
18. Add monitor on the forwarder by executing below line

```
./splunk add monitor /var/log/httpd -index main -sourcetype UFserverlogs
```

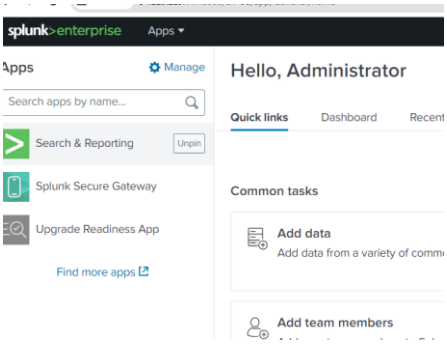
```
[root@ip-172-31-44-178 /]# /opt/splunkforwarder/bin/splunk add monitor /var/log
/httpd -index main -sourcetype UFserverlogs
Warning: Attempting to revert the SPLUNK_HOME ownership
Warning: Executing "chown -R splunkfwd:splunkfwd /opt/splunkforwarder"
Your session is invalid.  Please login.
Splunk username: satzsplunk
Password:
```

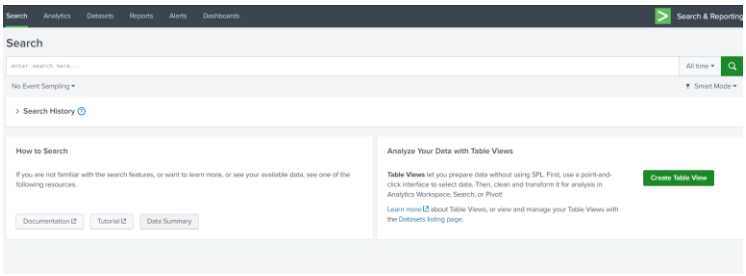
It adds monitoring for the "/var/log/httpd " directory, associates it with the "main" index, and assigns the sourcetype "UFserverlogs" for log data. This helps Splunk collect and index data from that directory with specified configurations.

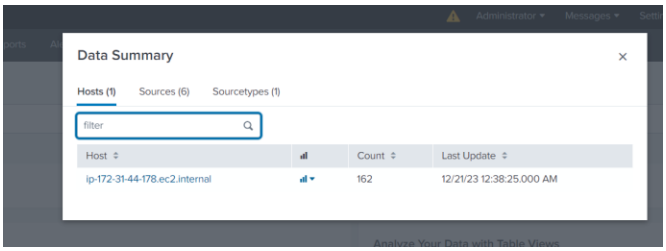
19. Refresh you App web page couple of time to generate some logs



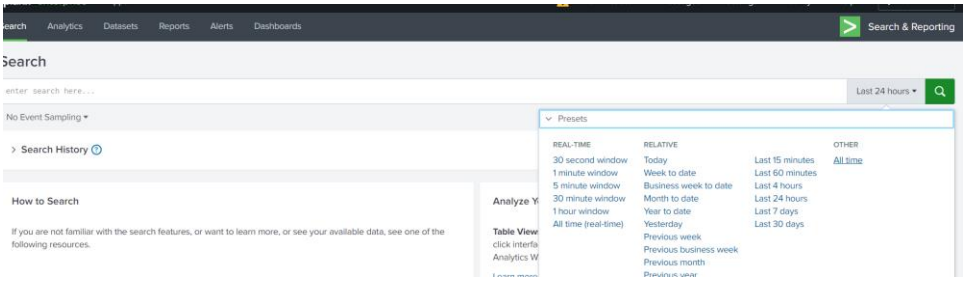
20. Go to Splunk site you hosted, click on Search and reporting, click on data summary, select host.

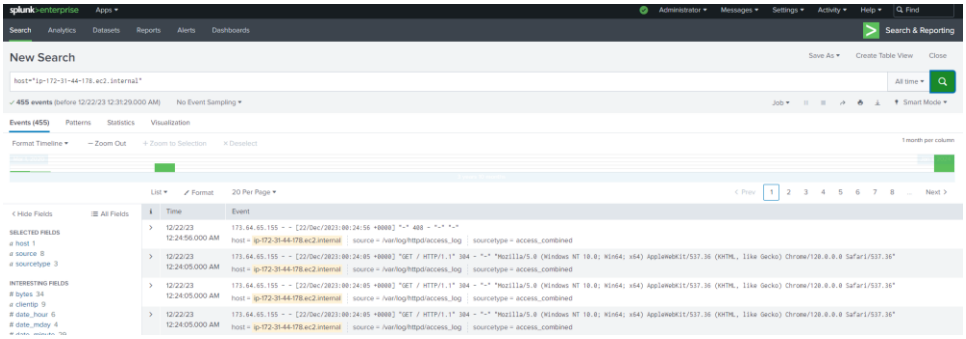






21. Click Last 24 hours and select All Time, to see all time log and click search, this will show the logs from the other machine.





Troubleshoot: In case you are not getting latest data from the forwarder; it may be your Splunk process don't have access to directories `/var/log/httpd`. Give read permission to directories.

Some common splunk queries are,

```
sourcetype=access_combined status=404
```

```
sourcetype=access_combined error*
```

```
sourcetype=access_combined clientip="192.168.1.100"
```

These examples provide a glimpse into the power of Splunk. Explore its extensive capabilities to address specific needs within your enterprise.

1. Security Analytics:
Identifying brute-force attacks:
```
sourcetype=auth* status!=200 | stats count by user, status | where count > 10
```

Detecting potential SQL injection attempts:
```
sourcetype=access_combined url=*.php* | search error OR (status=400 OR status=500) | regex "select|insert|update|delete"
```

Tracking anomalous user behavior:
```
sourcetype=wineventlog EventCode=4624 Logon_Type=3 | stats count by user | baseline count over last 7 days by user | where current > baseline * 3
```

2. Operational Monitoring:
Identifying server performance bottlenecks:
```
sourcetype=access_combined response_time > 5 | timechart span=1m avg(response_time)
```

Tracking application error trends:
```
sourcetype=error_log | timechart span=1h count by application
```

Monitoring resource usage:
```
sourcetype=cpu OR sourcetype=memory | timechart span=5m avg(pct_usage) by host
```

3. Business Analytics:
Analyzing customer behavior:
```
sourcetype=access_combined | search action=purchase | timechart span=1d count by product_id
```

Measuring website performance:
```
sourcetype=access_combined | stats avg(response_time) by url | sort avg(response_time)
```

Tracking sales trends:
```
sourcetype=sales_data | timechart span=1w sum(revenue) by region
```

4. Advanced Techniques:
Using subsearches and field lookups:
```
index=web sourcetype=access_combined | lookup user_info user as user_id OUTPUT username | search username=admin
```

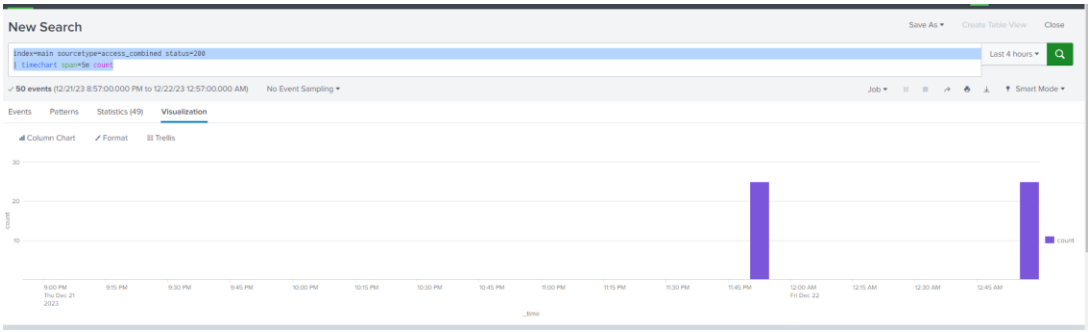
Joining multiple data sources:
```
index=web sourcetype=access_combined | join type=inner [search index=app sourcetype=app_logs] on user_id
```

Applying statistical functions and machine learning:
```
sourcetype=access_combined | anomalydetection algo=3sigma by user_id | table user_id, anomalous_value
```

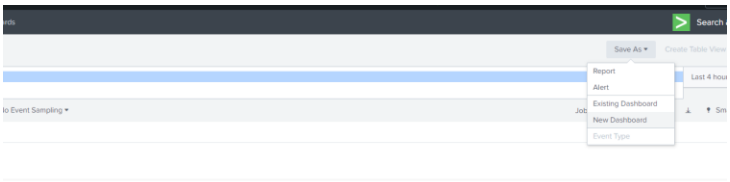
Splunk Dashboard

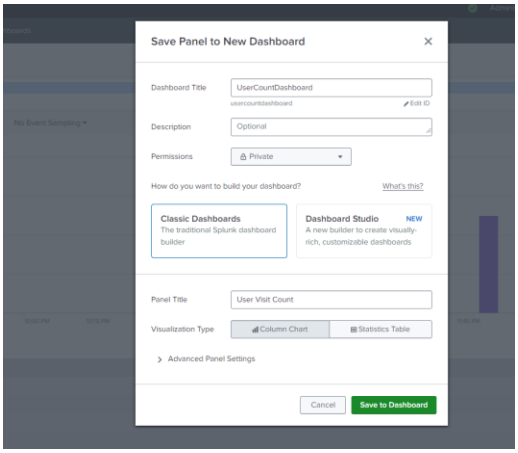
Let's create a simple Dashboard with the below query to see the users access attempt on the webpage for every 5 min frequency.

```
index=main sourcetype=access_combined status=200
| timechart span=5m count
```



Add query and click on Visualization to see values in graph. Select SaveAs and create a New Dashboard.





Now the dashboard is ready to view and can pin it to the home page.